THE ASTROPHYSICAL JOURNAL, 518:L119–L122, 1999 June 20
© 1999. The American Astronomical Society. All rights reserved. Printed in U.S.A.

OVERSHOOTING IN MASSIVE STARS: CONCEPTUAL PROBLEMS AND SUGGESTED SOLUTION

V. M. CANUTO
NASA Goddard Institute for Space Studies, 2880 Broadway, New York, NY 10025; and Department of Applied Physics, Columbia University, New York, NY 10027; vcanuto@giss.nasa.gov
Received 1999 March 22; accepted 1999 April 16; published 1999 May 4

ABSTRACT

Overshooting (OV) in massive stars faces a basic difficulty: *what theoretical models offer is not what stellar structure studies need*. The former use $\nabla_\mu = 0$ and define the OV where the negative convective flux J vanishes, while the latter need to know where the concentration flux J_c vanishes. We suggest that J may be dynamically irrelevant and derive the new dynamic equations for J_c. A new feature emerges: for large concentration gradients ∇C, J_c is no longer proportional to ∇C, as always assumed.

Subject headings: convection — stars: evolution — stars: interiors

1. THE PROBLEM

For the overshooting (OV) problem, it is the best of times and it is the worst of times. It is the best of times because theoretical models exist to quantify the OV based on the behavior of the convective flux J. It is the worst of times because the OV so obtained is not what stellar structure studies call OV. The mismatch between supply and demand has forced "users" to either disregard the theoretical predictions, parameterize the OV mixing empirically and obtain a result by fitting the data, or use the theoretical criteria only to conclude that what they need is different from what the models predict.

The reason for the chasm is as follows. On the supply side, theorists define the OV as the extent within the stably stratified zone (outside the convective zone [CZ]) where the convective flux is negative and erudite discussions can be found as to how long ∇ can remain close to $\nabla_{\rm ad}$ (like in the CZ) before detaching from it and becoming ∇_r (the radiative delta). Even without explicit calculations, it is fairly clear that as one moves away from the CZ, ∇ is no longer $\sim\nabla_{\rm ad}$ since strong mixing is no longer available and the eddies must now survive on a much leaner energy supply coming from nonlocal effects that play the role of a tenuous "source." The conclusion is that such a region must be very small indeed.

On the demand side, stellar structure studies of evolved convective cores have little use for such results. What they need to know is how far mixing occurs: material of the main CZ with a mean molecular weight μ_1 is "overshot" into the stably stratified region where the average μ_2 is lower than μ_1. The net result is that this region ends up with a μ higher than before the intrusion. This affects the position of the star in the H-R diagram.

Let us consider the two fluxes, *convective flux* J and *concentration flux* J_c,

$$J(r) \equiv \overline{w'T'}, \quad J_c(r) \equiv \overline{w'c'}. \tag{1a}$$

Theoretical models (see, e.g., Roxburgh 1978; Canuto 1997) quantify the OV using J while stellar structure studies use J_c. The first enters the mean temperature equation while J_c enters the mean concentration equation (χ_c is the kinematic diffusivity of c):

$$\rho \frac{DC}{Dt} = (\rho\chi_c C_{,i})_{,i} - (\rho J_i^c)_{,i}. \tag{1b}$$

J and J_c can be related using the relation $\rho'/\rho = -\alpha_T T' + \alpha_c c'$, where $\alpha_T \equiv -\rho^{-1}\partial\rho/\partial T$ and $\alpha_c \equiv \rho^{-1}\partial\rho/\partial C$ are the volume expansion coefficients. We derive the following relation:

$$J_\rho = -\alpha_T J + \alpha_c J_c, \quad J_\rho = \rho^{-1}\overline{w'\rho'}. \tag{1c}$$

The role of the *mass flux* J_ρ can be seen by noticing that the dynamic equation for the turbulent kinetic energy K does not depend on J and/or J_c separately but only on J_ρ:

$$\frac{\partial K}{\partial t} + \frac{\partial}{\partial z}F_{\kappa e} = -gJ_\rho - \epsilon, \quad F_{\kappa e} \equiv \tfrac{1}{2}\overline{w'q^2}, \tag{1d}$$

where $F_{\kappa e}$ is the flux of turbulent kinetic energy $K = \frac{1}{2}\overline{q^2}$. A further insight can be gained by considering the density equation. Separating mean and fluctuating parts and averaging, one obtains the following result for the mean density (h denotes the horizontal component):

$$\mathrm{v}_h \cdot \nabla_h\rho + w\frac{\partial}{\partial z}\rho = -\frac{\partial}{\partial z}(\rho J_\rho). \tag{1e}$$

In the one-dimensional model, the fluid velocity w is given by (N is the Brunt-Väisälä frequency)

$$w = g(\rho N^2)^{-1}\frac{\partial}{\partial z}(\rho J_\rho),$$

$$N^2 = -g\rho^{-1}\frac{\partial\rho}{\partial z} = gH_p^{-1}(\nabla_\mu - \nabla + \nabla_{\rm ad}). \tag{1f}$$

The macroscopic velocity w *is governed by the mass flux* J_ρ*. If* J *is dynamically irrelevant, as we discuss next,* w *is governed by* J_c.

2. THE DYNAMICAL ROLE OF THE CONVECTIVE FLUX J

Here we discuss the possibility that J may be very small in the stably stratified region outside the CZ and thus dynamically irrelevant. Numerical simulations (Freytag, Ludwig, & Steffen 1996) confirm this conjecture but do not explain why this is so. Let us call $E(k)$, $E_\theta(k)$, ϵ, and ϵ_θ the spectra of $\overline{w'^2}$ and $\overline{T'^2}$ and their corresponding rates of dissipation by viscosity and thermal conductivity. Stable stratification makes the flow

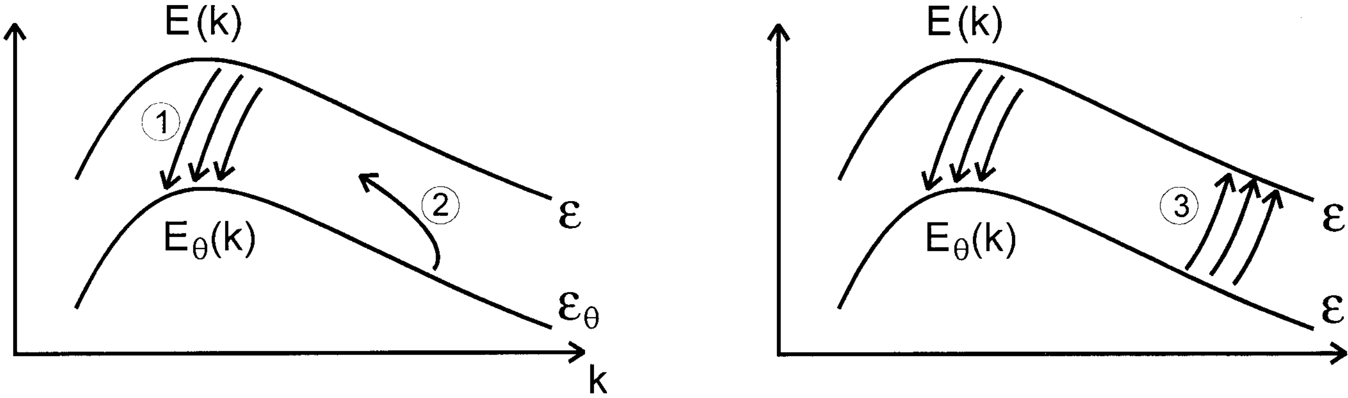

FIG. 1.—*Left*: In a stably stratified regime, kinetic energy transforms into potential energy (process 1). To obtain $\epsilon \approx \epsilon_\theta$, Weinstock (1981) suggested that some potential energy is backscattered (process 2). *Right*: An alternative possibility (Holloway 1988) is that at some k, potential energy transforms back into kinetic energy (process 3).

behave two-dimensionally: since the energy flow is uphill, the turbulent kinetic energy (TKE) of the large scales cannot cascade toward smaller scales, and the only route is to change into potential energy (PE), $\overline{T'^2}$ (process 1, left panel of Fig. 1). This is accompanied by a negative temperature flux $J < 0$. Contrary to TKE, PE can cascade to smaller scales; this results in little TKE but large PE, which is dissipated, and thus $\epsilon_\theta \gg \epsilon$. Available data do not confirm this result, but rather they point to $\epsilon_\theta \approx \epsilon$. Thus, some PE must somehow be prevented from cascading all the way to the smaller scales. There are two possibilities: Weinstock (1981) suggested that some PE is backscattered to larger scales (process 2, left panel of Fig. 1). This reduces the amount of PE to be dissipated and leads to a more equal ϵ and ϵ_θ, as observed. The problem with this scenario is that it does not clarify the fate of the PE now accumulated at the largest scales, which can be tapped by many instabilities. In that respect, the model is incomplete. The other suggestion (Holloway 1988) is that at some intermediate k, PE is transformed back to TKE (process 3, right panel of Fig. 1), thus alleviating ϵ_θ while enriching ϵ. If so, smaller scales would become *unstably stratified* with a positive J. In this scenario, the $J(k)$ spectrum looks like that in Figure 2, and since the positive portion tends to compensate the negative one, one ends up with a small negative J that in turn implies a small J-based OV, as indeed is found. Recent numerical simulations confirm the spectrum in Figure 2 (Bouruet-Aubertot, Sommeria, & Staquet 1996). Thus, if J is dynamically irrelevant in the OV region, one can understand why helioseismology yields a small J-based OV. This does not imply, however, that the mixing-based OV is also small.

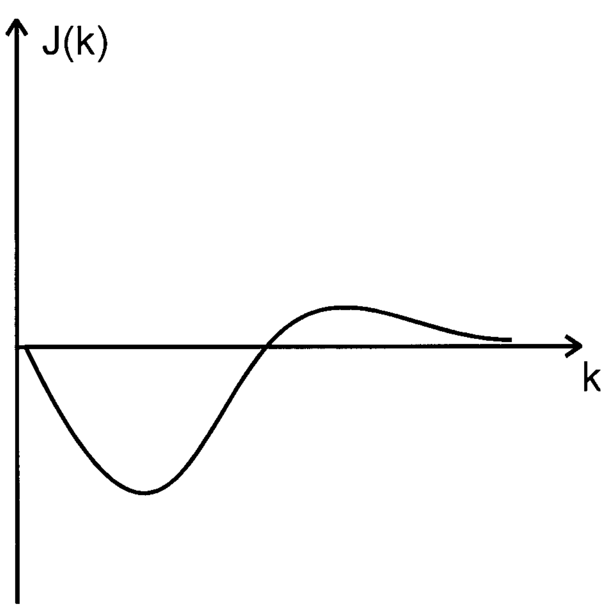

FIG. 2.—In the case of process 3 in the right panel of Fig. 1, the convective flux J is partly positive and partly negative. This results in a small J (Holloway 1988; Bouruet-Aubertot et al. 1996).

3. EQUATIONS FOR J AND J_c

Deriving the dynamic equations for J and J_c is not easy, and until recently such equations were not available. Canuto & Dubovikov (1998, hereafter CD) derived the dynamic equations governing the evolution of turbulence with $\nabla_\mu = 0$. Before extending such equations to the $\nabla_\mu \neq 0$ case, one needs some assurance that the model yields reliable results. The CD model is a particular case for convection of a more general model for arbitrary turbulent flows (Canuto & Dubovikov 1996a, 1996b, 1996c, 1997a, 1997b, 1997c, 1999) that has thus far reproduced some 80 statistics corresponding to a wide range of turbulent flows. Furthermore, Kupka (1999) has shown that the CD model compares well with numerical simulations.

The extension of CD to the $\nabla_\mu \neq 0$ case has recently been presented (Canuto 1999). The turbulent part of the velocity-temperature fields is governed by the five variables:

$$K \equiv \tfrac{1}{2}\overline{u_i'u_i'}, \quad K_z \equiv \tfrac{1}{2}\overline{w'^2}, \quad \mathrm{PE} \equiv \tfrac{1}{2}\overline{T'^2}, \quad J \equiv \overline{w'T'}, \quad \epsilon, \tag{2a}$$

which represent the total turbulent kinetic energy, the z-component of it, the potential energy, the convective flux, and ϵ (the rate of dissipation of K). The dynamic equations are as follows:

$$\frac{\partial K}{\partial t} + D_f = -gJ_\rho - \epsilon, \quad (3a)$$

$$\frac{\partial}{\partial t} K_z + D_f = -5\tau^{-1}\left(K_z - \frac{1}{3}K\right) - \frac{2}{3} gJ_\rho - \frac{1}{3}\epsilon, \quad (3b)$$

$$\frac{\partial}{\partial t} J + D_f = 2\beta K_z + (1-\gamma_1)g\alpha_T\overline{T'^2} - \tau_{p\theta}^{-1}J + \tfrac{1}{2}\frac{\partial}{\partial z}\left(\chi \frac{\partial}{\partial z} J\right) + F_c, \quad (3c)$$

$$\frac{\partial}{\partial t}\tfrac{1}{2}\overline{T'^2} + D_f = \beta J - \tau_\theta^{-1}\overline{T'^2} + \tfrac{1}{2}\frac{\partial}{\partial z}\left(\chi \frac{\partial}{\partial z}\overline{T'^2}\right), \quad (3d)$$

$$\frac{\partial \epsilon}{\partial t} + D_f = -c_1 gJ_\rho \epsilon K^{-1} - c_2\epsilon^2 K^{-1}, \quad (3e)$$

where

$$F_c \equiv -g\alpha_c\overline{T'c'} + \tfrac{1}{2}\frac{\partial}{\partial z}\left(\chi_c \frac{\partial}{\partial z} J_c\right). \quad (3f)$$

When $J_\rho \to -\alpha_T J$ and $F(J_c) = 0$, equations (3a)–(3e) coincide with those solved by Kupka (1999). The only equation not affected by the c-field is the one for the potential energy. In equations (3a)–(3f), the nonlocal terms, D_f, the relation of the timescales to the dynamical timescale, $\tau \equiv 2K\epsilon^{-1}$, and the constants are discussed and given in CD. Once we include the c-field, three new turbulent correlations appear:

$$\overline{w'c'}, \quad \tfrac{1}{2}\overline{c'^2}, \quad \overline{T'c'}, \quad (4a)$$

representing the concentration flux J_c, the potential energy of the c-field, and the c-T correlation. The dynamic equations for the variables (4a) are as follows:

$$\frac{\partial}{\partial t} J_c + D_f = -2K_z\frac{\partial C}{\partial z} - g\alpha_c\overline{c'^2} - \tau_{pc}^{-1}J_c + g\alpha_T\overline{T'c'} + \tfrac{1}{2}\chi_c\frac{\partial^2}{\partial z^2}J_c, \quad (4b)$$

$$\frac{\partial}{\partial t}\tfrac{1}{2}\overline{c'^2} + D_f = -J_c\frac{\partial C}{\partial z} - \tau_\theta^{-1}\overline{c'^2}, \quad (4c)$$

$$\frac{\partial}{\partial t}\overline{c'T'} + D_f = \beta J_c - J\frac{\partial C}{\partial z} - \frac{15}{2}\sigma_t^{-1}\tau^{-1}\overline{c'T'}. \quad (4d)$$

There are eight turbulent variables (eqs. [2a] and [4a]) and eight coupled dynamic equations; $\beta = TH_p^{-1}(\nabla - \nabla_{\rm ad})$, $\sigma_t = 0.72$ is the turbulent Prandtl number, and

$$\alpha_c\frac{\partial C}{\partial z} = -H_p^{-1}\nabla_\mu. \quad (4e)$$

The mean temperature equation is unchanged while the mean concentration equation is given by equation (1b). If a stellar model provides $\beta(r)$, $g(r)$, $\chi(r)$, and $\alpha(r)$, equations (3a)–(4e) can be solved and the μ-based OV determined by the behavior of the concentration flux $J_c(r)$.

4. SOLUTION OF THE CONCENTRATION EQUATIONS

Here we consider the local limit of equations (4b)–(4d). Neglecting the nonlocal terms D_f and taking the stationary case, equations (4b)–(4d) give

$$J_c = -D_c\frac{\partial C}{\partial z}, \quad D_c = \frac{d}{1-\eta}, \quad (5a)$$

where D_c is the "*turbulent concentration diffusivity*" with

$$\mathrm{d} = \frac{2}{5}\sigma_t(1+\sigma_t)^{-1}\left(\tau K_z + \frac{1}{15}\sigma_t\alpha_T g\tau^2 J\right), \quad (5b)$$

$$\eta = \frac{1}{5}\sigma_t^2(1+\sigma_t)^{-1}\left(\frac{2}{15}g\alpha_T\beta\tau^2 + \tau^2 g\alpha_c\frac{\partial C}{\partial z}\right). \quad (5c)$$

Only if we neglect all the terms except the one with K_z, and only if we take the local limit of equation (3e), that is, $\epsilon = K^{3/2}\Lambda^{-1}$, where Λ is the mixing length, would we get a diffusivity of the form used thus far in all calculations, namely, $D_c \sim \Lambda w$, while equations (5a)–(5c) show a considerably more complex structure. D_c depends also on the temperature gradient β, on the convective flux J, and *on the concentration gradient itself*, which is a new feature not exhibited by any previous model. Finally, the model contains no adjustable parameters and no mixing length.

5. A NEW FEATURE: THE CASE OF A LARGE $\partial C/\partial z$

Equations (5a)–(5c) exhibit a new feature: for large $\partial C/\partial z$, the use of $\tau = 2K\epsilon^{-1}$ and $3K_z = K$ gives

$$J_c = (3\sigma_t g\alpha_c)^{-1}\epsilon. \quad (6a)$$

The concentration flux no longer depends on the concentration gradient as always assumed.

6. AN ILLUSTRATIVE SOLUTION IN THE CASE OF LARGE $\partial C/\partial z$

In the stationary limit, equations (3a) and (3e) become

$$\frac{\partial}{\partial z}F_{\kappa e} = -\left(1+\frac{1}{3\sigma_t}\right)\epsilon,$$

$$\frac{\partial}{\partial z}F_{\kappa\epsilon} = -\left(\frac{c_1}{3\sigma_t}+c_2\right)\epsilon^2K^{-1}. \quad (6b)$$

For the fluxes $F_{\kappa e}$ and $F_{\kappa\epsilon}$ of K and ϵ, we use the closures

$$F_{\kappa e} = -\nu_t\frac{\partial K}{\partial z}, \quad F_{\kappa\epsilon} = -\nu_t\sigma_t^{-1}\frac{\partial\epsilon}{\partial z},$$

$$\nu_t = C_\mu K^2\epsilon^{-1}, \quad (6c)$$

which have been widely used in turbulence studies with $C_\mu \approx 0.08$; ν_t plays the role of turbulent viscosity. The Fickian closures (eqs. [6c]) are not the best one can construct, and more physical ones have been constructed (Canuto 1993), but in the interest of carrying out an analytical model, the representation in equations (6c) has a sufficient amount of physics to be worthwhile considering. Kupka (1999) has used both the first expression in equations (6c) and the more complete model for $F_{\kappa e}$ and has shown that the latter gives a much better fit to the data, which is something we can at least qualitatively achieve by boosting C_μ from its nominal value. Interestingly enough, the two coupled differential equations (eqs. [6b] and [6c]) can be integrated analytically. We only outline the necessary steps: first, multiply both sides of both equations (6b) by ν_t and introduce the new variable, $\nu_t\partial/\partial z = \partial/\partial\xi$. Equations (6b) and (6c) have the solutions $K(\xi) = A\xi^{-2}$ and $\epsilon(\xi) = B\xi^{-2}$. The condition $(3\sigma_t)^{-1} + 1 = c_2\sigma_t + c_1/3$ must be satisfied. Returning to the variable z, one obtains

$$K(z) = K_0(1 - z/l_k)^2, \quad \epsilon(z) = \epsilon_0(1 - z/l_\epsilon)^2, \quad (6d)$$

where the scales $l_{\kappa,\epsilon}$ are defined by

$$l_k \equiv C_\mu K_0^{1/2}A^{3/2}B, \quad l_\epsilon \equiv l_k(A\epsilon_0/BK_0)^{1/2},$$

$$A \equiv 6C_\mu^{-1}\left(1+\frac{1}{3\sigma_t}\right)^{-1}. \quad (6e)$$

Thus, because of equation (6a), *the concentration flux J_c vanishes where $\epsilon(z)$ vanishes, and that is the endpoint of the OV region.* The constant B cannot be fixed because the equation for ϵ becomes linear in ϵ. However, if we demand that $l_{k,\epsilon} \equiv l$, then

$$l = \left(\frac{18\sigma_t C_\mu}{1+3\sigma_t}\right)^{1/2}\frac{K_0^{3/2}}{\epsilon_0} = 2C_\mu^{1/2}\frac{K_0^{3/2}}{\epsilon_0}, \quad (6f)$$

where $\sigma_t = 0.72$. The values K_0 and ϵ_0 can be taken at the beginning of the stably stratified region. As discussed before, one probably has to take a C_μ larger than its nominal value, and if so, equation (6f) implies a larger l. It may then become justified to approximate equations (6d) by

$$K(z) = K_0 \exp(-2z/l), \quad (6g)$$

a behavior suggested by numerical simulations (Freytag et al. 1996). There are two more results of interest. Substituting equation (6a) into equations (1f), we obtain, with $J_\rho \approx \alpha_c J_c$,

$$w(z) = (3\sigma_t\rho N^2)^{-1}\frac{\partial}{\partial z}(\rho\epsilon), \quad (6h)$$

which shows how the velocity decreases with z. Finally, if we use the closure

$$\rho J_\rho = -K_\rho\frac{\partial\rho}{\partial z}, \quad (7a)$$

where K_ρ is the *mass diffusion coefficient*. Equation (7a) yields the relation

$$K_\rho = \gamma\epsilon N^{-2}, \quad \gamma = (3\sigma_t)^{-1}, \quad (7b)$$

which has been widely used in oceanography (Gargett 1989).

7. CONCLUSIONS

For many years, people tried to determine the OV using the convective flux $J(r)$, an inadequate variable for it is dynamically irrelevant. It must be replaced with the concentration flux $J_c(r)$ for which we have derived the dynamic equations. A new feature has emerged: for large concentration gradients ∇C, the concentration flux is no longer proportional to ∇C, as usually assumed. We have presented an illustrative analytic solution so as to exhibit the extent of the OV as determined from the behavior of the concentration flux. The next step is to solve equations (3a)–(3f) and equations (5a)–(5c) in conjunction with a stellar model. In conclusion, the OV extent in massive stars is determined by $J_c(r)$ and not by $J(r)$.

REFERENCES

Bouruet-Aubertot, P., Sommeria, J., & Staquet, C. 1996, Dyn. Atmos. Oceans, 23, 357
Canuto, V. M. 1993, ApJ, 416, 331
———. 1997, ApJ, 489, L71
———. 1999, ApJ, in press
Canuto, V. M., & Dubovikov, M. S. 1996a, Phys. Fluids, 8, 571
———. 1996b, Phys. Fluids, 8, 587
———. 1996c, Phys. Fluids, 8, 599
———. 1997a, Phys. Fluids, 9, 2118
———. 1997b, Phys. Fluids, 9, 2132
Canuto, V. M., & Dubovikov, M. S. 1997c, Phys. Fluids, 9, 2141
———. 1998, ApJ, 493, 834 (CD)
———. 1999, Phys. Fluids, 11, 659
Freytag, B., Ludwig, H. G., & Steffen, M. 1996, A&A, 313, 497
Gargett, A. E. 1989, Annu. Rev. Fluid Mech., 21, 419
Holloway, G. 1988, Dyn. Atmos. Oceans, 12, 107
Kupka, F. 1999, in Proc. Granada Workshop on Stellar Convection, in press
Roxburgh, I. 1978, A&A, 65, 281
Weinstock, J. 1981, J. Atmos. Sci., 38, 880